

Thematische Abfrage mit Computerlinguistik

Autor: Dr. Klaus Loth (ETH-Bibliothek Zürich)

Zusammenfassung

Der Beitrag befasst sich mit dem Einsatz der Computerlinguistik bei der thematischen Abfrage einer mehrsprachigen bibliographischen Datenbank. Das Verbundsystem NEBIS (Netzwerk von Bibliotheken und Informationsstellen in der Schweiz) wurde durch computerlinguistische Komponenten ergänzt, um die thematische Abfrage in den drei Sprachen Deutsch, Englisch und Französisch effektiver und benutzerfreundlicher zu machen. Im Schwerpunkt wird die Kopplung des dreisprachigen Sachregisters (Thesaurus) mit dem computerlinguistischen System behandelt.

Thematische Abfrage mit Computerlinguistik

Die ETH-Bibliothek und ihre Verbundpartner sammeln vorwiegend Dokumente in deutscher, englischer und französischer Sprache. Die ETH-Bibliothek und ihre klassifizierenden Verbundpartner verwenden zur inhaltlichen Beschreibung der Dokumente ein dreisprachiges (Deutsch, Englisch, Französisch), hierarchisch strukturiertes Sachregister (Thesaurus). Im Verbund gibt es allerdings auch Partner, die deutsche Schlagwörter zur inhaltlichen Beschreibung verwenden. Der Verbund besteht zur Zeit aus 85 Bibliotheken und weist einen Bestand von etwa 3 Millionen Titeln auf.

Mit Hilfe der Computerlinguistik können diese reichlich unterschiedlichen Bedingungen auf einen vernünftigen gemeinsamen Nenner gebracht werden.

Folgende Daten werden der computerlinguistischen Analyse unterworfen.

Titel, Untertitel,
Begriffe aus dem dreisprachigen Sachregister (Thesaurus)
und deutsche Schlagwörter

Für die thematische Recherche sieht das System natürlichsprachliche Eingaben im Sinne der Schriftsprache vor. Der Benutzer kann also bei seiner Datenbankrecherche die Suchanfrage in dieser Weise formulieren. Auf der Basis von computerlinguistischen Analysen wird die Suchanfrage auf der Datenbank syntaktisch-semantisch prozessiert. Ein speziell entwickelter Ranking-Algorithmus gewichtet, bewertet und relationiert Suchergebnisse nach Themen, Einschränkungen und Modifikationen. Die erkannten syntaktisch-semantischen Zusammenhänge gehen dabei nicht durch die 'sinnleerende' Reduktion auf Boolesche Verknüpfungen verloren.

Eine ausführliche Beschreibung finden Sie unter
Loth, Klaus: Thematische Abfrage einer dreisprachigen Datenbank mit computerlinguistischen Komponenten; ABI Technik 24 (2004) 4, p. 294 – 300.

Im Folgenden soll die Kopplung des Thesaurus mit dem computerlinguistischen System an Hand einiger Beispiele gezeigt werden.

Der Benutzer schreibt sein gewünschtes Thema in das Eingabefeld.

Stichwortsuche:

Stichwörter aus allen Feldern
Titel enthält Stichwörter...
Autor/Körperschaft enthält Stichwörter...
Serie/Reihe enthält Stichwörter...

Beispiele:

frisch andorra
frisch and andorra
anthropolog? china
(herz or cardio?) and therap?

Blättern in einer Liste:

Autor/Körperschaft beginnt mit...
Titel beginnt mit...
Zeitschriftentitel beginnt mit...
Sachbegriff NEBIS (D,E,F) beginnt mit...

Beispiele:

frisch m
frisch max
schweizerische afrika gesellschaft
conference on acid rain

Tipp: bei Autoren Nachname zuerst

Themen - natürlichsprachliche Abfrage:

Suche Literatur zum Thema ...

Beispiele:

einführung in die psychologie
wissenschaft aus historischer sicht
neutron scattering
c*-algebras

[Info](#)

Das System sucht gleichzeitig die Titel zum Thema und im Thesaurus die Einträge, in denen der Begriff „Organische Chemie“ oder Derivate davon vorkommen.

Zuerst werden die Titel mit dem entsprechenden Ranking angezeigt.

Themen - natürlichsprachliche Abfrage (TEST):

Suche Literatur zum Thema ...

Suchergebnis filtern nach

Dokumentenart:

alle

Erscheinungsjahr:

=
(1111)

Suchergebnis erweitern um

Abstract & Index:

ohne

Beispiele:

einführung in die psychologie
wissenschaft aus historischer sicht
neutron scattering
c*-algebras

Ergebnisse für Wörter=(Organische Chemie)

Treffer: > 500 (~ 18600) (3.98 Sek.)

#.	Autor	Titel	Jahr
1. [100%]	Breitmaier, Eberhard	Organische Chemie : Grundlagen, Stoffklassen, Reaktionen, Konzepte, Molekülstruktur	2001
2. [100%]	Fodor, Gábor	Organische Chemie : mit Literatur- und Registerband	1965
3. [100%]	Asselborn, Wolfgang	Chemie heute	2001
4. [100%]	Hückel, Walter	Organische Chemie	1947
5. [100%]	Asselborn, Wolfgang	Organische Chemie	1991
6. [100%]	Latscha, Hans Peter	Organische Chemie	2002
7. [100%]	Hart, Harold	Organische Chemie	2002

Am Ende dieser Seite befindet sich folgende Anzeige:

Siehe auch:

[Thematische Begriffe \(ETH Thesaurus\)](#)








Wenn im Thesaurus zu viele Begriffe mit „organische Chemie“ oder Derivate davon gefunden werden, wird die Anzeige „Thematische Begriffe (ETH Thesaurus)“ gemacht.

Wird auf „Thematische Begriffe“ geklickt, folgt eine Liste der entsprechenden Begriffe aus dem Thesaurus.

Thematische Begriffe - natürlichsprachliche Abfrage:	Beispiele: Frauen und Wissenschaft Lineare Algebra Neutronenstreuung C*-Algebren
Suche im Thesaurus Informationen zum Thema ... <input type="text" value="Organische Chemie"/> <input type="button" value="OK"/>	

Ergebnisse für Wörter=(Organische Chemie)

Treffer: 25 (0.01 Sek.)






#. [Rel. %]	Thematische Begriffe	Titel
1. [100%]	ORGANISCHE CHEMIE	
2. [98%]	BIOANORGANISCHE CHEMIE, ANORGANISCHE SUBSTANZEN UND MINERALSTOFFE	
3. [97%]	MAKROMOLEKÜLE/ORGANISCHE CHEMIE	
4. [97%]	PHYSIKALISCHE ORGANISCHE CHEMIE	
5. [97%]	THEORETISCHE ORGANISCHE CHEMIE	
6. [97%]	ELEKTROORGANISCHE CHEMIE	
7. [90%]	OXOCARBONSÄUREN (ORGANISCHE CHEMIE)	

Wird nun auf 1. geklickt, folgt die Anzeige des Thesaurus an der Stelle „Organische Chemie“.

Relationen	Thematische Begriffe	Titel
Oberbegriff	CHEMIE	
Begriff	ORGANISCHE CHEMIE	
Unterbegriff	PHYSIKALISCHE ORGANISCHE CHEMIE	
Unterbegriff	ORGANISCH-CHEMISCHE ANALYSE	
Unterbegriff	THEORETISCHE ORGANISCHE CHEMIE	
Unterbegriff	ORGANISCHE SYNTHESE (CHEMIE)	
Unterbegriff	ORGANISCHE IONEN	
Unterbegriff	ORGANISCHE FESTKÖRPERCHEMIE	
Unterbegriff	ORGANISCHE VERBINDUNGEN U. ORGANISCHE MOLEKÜLE	
Unterbegriff	ORGANISCH-CHEMISCHE REAKTIONEN + ORGANISCH-CHEMISCHE REAKTIONSMechanismen	

Die Benutzer haben nun die Möglichkeiten, entweder die Titel zu einem bestimmten Begriff anzusehen, indem sie auf das entsprechende gelbe Ordnersymbol klicken oder den Thesaurus weiter zu inspizieren, indem sie auf den gewünschten Begriff klicken.

Klick auf „Organische Ionen“.

Relationen	Thematische Begriffe	Titel
Oberbegriff	<u>ORGANISCHE CHEMIE</u>	
Oberbegriff	IONISierter Zustand (CHEMIE)	
Begriff	ORGANISCHE IONEN	
Unterbegriff	ORGANISCHE KATIONEN	
Unterbegriff	ORGANISCHE ANIONEN	
Unterbegriff	NITRENIUMIONEN + NITRENE (ORGANISCHE CHEMIE)	

Auf diese Weise können die Benutzer auf eine einfache Art die Titel finden und gleichzeitig werden sie über die Begriffsstruktur des Thesaurus informiert.

Es ist ja denkbar, dass bei der Suche nach dem Thema ein zu grober Begriff benutzt wurde. Bei der Eingabe weiss ein Benutzer natürlich noch nicht, welche Begriffe verwendet werden und wie diese formuliert sind. Der Thesaurus informiert nun die Benutzer über die möglichen Begriffe, die für die thematische Suche zur Verfügung stehen.

Nehmen wir an, dass eigentlich Informationen zu „organischen Kationen“ gesucht wurden, so wird der Benutzer nun auf das Ordnersymbol hinter diesem Begriff klicken.

Themen - natürlichsprachliche Abfrage (TEST):

Suche Literatur zum Thema ... **Beispiele:**
 einföhrung in die psychologie
 wissenschaft aus historischer sicht
 neutron scattering
 c*-algebras

Suchergebnis filtern nach Dokumentenart:

Erscheinungsjahr:

Suchergebnis erweitern um Abstract & Index:

Ergebnisse für Wörter=(ORGANISCHE KATIONEN) Filter: erweitert um Abstract & Index

Treffer: 78 (0.00 Sek.)

[◀ Vorherige Seite](#) [Nächste Seite ▶](#)

#. [Rel. %]	Autor	Titel	Jahr
1. [100%]	Thal, Arvid	Heterodimere aus anionischen und kationischen Porphyrinen mit variablem Abstand	2003
2. [100%]	Rinderhagen, Heiko	Cyclopropylsilyltherradikalkationen : photochemische Erzeugung, mechanistische Studien und synthetische Anwendung	2002
3. [100%]	Loss, Sandra	Niedervalente kationische Organophosphorverbindungen und phosphorzentrierte Radikale	2001
4. [100%]	Gotta, Matthias	Reaktionen kationischer Elektrophile mit neutralen Nucleophilen : Anwendungsbereich, Erweiterungsmöglichkeiten und Grenzen von Reaktivitätsskalen	2001
5. [100%]	Gillen, Ralph C.	Wellenpaketdynamik von konischen Durchschneidungen : Spektroskopie und ultraschneller Zerfall von trans-Butadien und Ethinkation	2000
6. [100%]	Franke, Holger	Die elektrophilen Kationen [C6F5N2] ⁺ und [C6F5Xe] ⁺ : Synthese und Reaktivität	2000
7. [100%]	Alickmann, Dirk	1-Aza- und 1-Oxa-polyenyalkationen	2000

In dieser Trefferliste werden alle 78 Titel angezeigt, die mit dem Begriff „organische Kationen“ klassifiziert wurden. Wie anfangs erwähnt, gibt es im Verbund aber auch Bibliotheken, die nicht mit den Thesaurusbegriffen klassifizieren. Um alle Titel zum Thema „organische Kationen“ in der gesamten Verbundsdatenbank zu finden, genügt ein Klick auf OK und alle Titel, auch die nicht sachbegrifflich so klassifizierten, werden gefunden, da im Eingabefeld „Suche Literatur zum Thema ..“ nun bereits „organische Kationen“ steht.

Ergebnisse für Wörter=(ORGANISCHE KATIONEN)

Treffer: > 500 (~ 1300) (2.21 Sek.)

[◀ Vorherige Seite](#) [Nächste Seite ▶](#)

#. [Rel. %]	Autor	Titel	Jahr
1. [94%]	Forrester, R.A.	Organische Kation-Radikale und Polyradikale	1980
2. [92%]	Auchter, Gerhard	Synthese und Solvolyse substituierter 1-Cyclobutenyl-Nonaflate	1983
3. [92%]	Bremer, Matthias	Experimentelle und theoretische Untersuchungen zur Adamantan- und Carbokationenchemie	1989
4. [92%]	Scheumann, Klaus	Dodecahedrane : Reaktionen von Carbokationen und Radikalen auf der Dodecahedran Oberfläche	1993
5. [90%]	Alickmann, Dirk	1-Aza- und 1-Oxa-polyenyalkationen	2000
6. [90%]	Gillen, Ralph C.	Wellenpaketdynamik von konischen Durchschneidungen : Spektroskopie und ultraschneller Zerfall von trans-Butadien und Ethinkation	2000
7. [90%]	Rinderhagen, Heiko	Cyclopropylsilyltherradikalkationen : photochemische Erzeugung, mechanistische Studien und synthetische Anwendung	2002

Jetzt hat das System mehr als 500 Titel gefunden.

Die bisherigen Beispiele wurden alle im deutschen Dialog gezeigt. Wird hier ein Titel zur Volltitelanzeige aufgerufen, so erscheint die Klassifikation in deutscher Sprache.

Vollansicht des Titels

Format wählen: [Standard Katalogkarte](#) [Zitierformat MARC](#)

Titel	Organische Kation-Radikale und Polyradikale : Substanzenverzeichnis für die Bände II/1 und II/9 = Organic cation radicals and polyradicals : Index of substances for volumes II/1 and II/9 / R. A. Forrester ... [et al.]
Impressum	Berlin [etc.] : Springer, 1980
Umfang	XII, 369 S. ; 28 cm : Ill.

Sacherschliessung	ORGANISCHE RADIKALE : 547.024*1
	RADIKALKATIONEN (CHEMIE) : 54.024-128.4
	FORMELTAFELN + ZAHLENTABELLEN : (083)
	FREIE RADIKALE (CHEMIE) : 541.515
	MAGNETISCHE EIGENSCHAFTEN : 537.621

Im System gibt es drei Dialogsprachen, Deutsch, Englisch und Französisch.

Die folgenden Beispiele werden im englischen Dialog gezeigt.

Topics - natural language search:
 Searching for literature about ...

Examples:
 Handbook of psychology
 A short history of technology
 neutron scattering
 c*-algebras

[Info](#)

Die Eingabe wird hier z.B. in deutscher Sprache gemacht. Nach Klick auf „Go“ werden wieder Titel zum Thema „Festkörperphysik“ gesucht.

Topics - natural language search:
 Searching for literature about ...

Filter search results by
 document type:
 year of publication: (yyyy)
 Expand search results with
 Abstract & Index:

Examples:
 Handbook of psychology
 A short history of technology
 neutron scattering
 c*-algebras

Results for words=(Festkörperphysik)

Hits: > 500 (~ 2500) (0.03 Sec.)

[Previous Page](#) [Next Page](#)

#.	Author	Title	Year
1. [100%]	Kittel, Charles	Einführung in die Festkörperphysik	1996
2. [100%]	Binkert, Thomas	Festkörperphysik : Vorlesung für Physikstudenten im 5. Semester (WS 1999/2000), sowie Ergänzungen und Übungen dazu (Festkörperphysik II)	2000
3. [100%]	Rice, Thomas Maurice	Festkörperphysik: Vergangenheit und Zukunft	1983
4. [100%]	Ibach, Harald	Solid-state physics : an introduction to principles of materials science	cop. 1995
5. [100%]	Busch, Georg	Beiträge zur Festkörperphysik : Prof. Dr. Georg Busch zu seinem 60. Geburtstag am 12. September 1968	
6. [100%]	Kittel, Charles	Einführung in die Festkörperphysik	2002

Am Ende dieses Bildschirmes steht

See also:



[PHYSICS OF CONDENSED MATTER](#)



[POROSITY AND PERMEABILITY \(MOLECUL.SYST.,PHYSICS\)](#)

Mit einem Klick auf „PHYSICS OF CONDENSED MATTER“ wird der Thesaurus an der Stelle „Festkörperphysik“ in englischer Sprache aufgerufen.

Relations	Thematical Terms	Titles
Broader Term	PHYSICS	
Term	CONDENSED MATTER MATTER	
Narrower Term	TRANSPORT PROCESSES, TRANSPORT PHENOMENA (CONDENSED-MATTER PHYSICS)	
Narrower Term	DIELECTRIC PROPERTIES (CONDENSED-MATTER PHYSICS)	
Narrower Term	MAGNETIC PROPERTIES (CONDENSED-MATTER PHYSICS)	
Narrower Term	SOFT CONDENSED MATTER	
Narrower Term	MESOSCOPIC SYSTEMS (CONDENSED-MATTER PHYSICS)	
Narrower Term	AMORPHOUS STATE OF SUBSTANCE (CHEMISTRY AND PHYSICS)	
Narrower Term	STRUCTURES + TRANSITIONS (CONDENSED-MATTER PHYSICS)	
Narrower Term	QUANTUM STATES (CONDENSED-MATTER PHYSICS)	
Narrower Term	OPTICAL PROPERTIES OF CONDENSED MATTER	
Related Term	STRUCTURE AND PROPERTIES OF MOLECULAR SYSTEMS (PHYSICS)	
Related Term	CRYSTALLOGRAPHY	

Mit einem Klick auf das Ordnersymbol hinter „SOFT CONDENSED MATTER“ werden die 30 Titel aufgerufen, die mit diesem Begriff, egal ob in deutscher, englischer oder französischer Sprache klassifiziert sind.

Topics - natural language search (TEST):

Searching for literature about ...

Filter search results by document type:

year of publication: (yyyy)

Expand search results with Abstract & Index:

Examples:
Handbook of psychology
A short history of technology
neutron scattering
c*-algebras

Results for words=(SOFT CONDENSED MATTER) Filter: expanded with Abstract & Index

Hits: 30 (0.01 Sec.)

[Previous Page](#) [Next Page](#)

#. [Rel. %]	Author	Title	Year
1. [100%]	Skjeltorp, Arne T.	Forces, growth, and form in soft condensed matter : at the interface between physics and biology	2004
2. [100%]	Popovi , Dragana	Fluctuations and noise in materials : 26-28 May 2004, Maspalomas, Gran Canaria, Spain	2004
3. [100%]	Gompper, Gerhard	Physics meets biology : from soft matter to cell biology	2004
4. [100%]	Case, Fiona	Mesoscale phenomena in fluid systems	2003
5. [100%]	Chabibullaev, Pulat Kirgizbaevi	Phase separation in soft matter physics : micellar solutions, microemulsions, critical phenomena	2003
6. [100%]	Kleman, Maurice	Soft matter physics : an introduction	2003

Mit einem weiteren Klick auf „Go“ werden alle Titel zu diesem Thema gesucht, auch die die nicht mit den Begriffen aus dem ETH-Thesaurus klassifiziert wurden.

Results for words=(SOFT CONDENSED MATTER)

Hits: 37 (0.23 Sec.)

[Previous Page](#) [Next Page](#)

#. [Rel. %]	Author	Title	Year
1. [100%]	Jones, Richard A.L.	Soft condensed matter	2002
2. [100%]	Helgesen, Geir	Soft condensed matter : configurations, dynamics and functionality	1999
3. [100%]	Skjeltorp, Arne T.	Soft condensed matter : configurations, dynamics and functionality	2000
4. [98%]	Lindner, Peter	Neutrons, X-rays and light : scattering methods applied to soft condensed matter	2002
5. [98%]	Riste, Tormod	Phase transitions in soft condensed matter	cop. 1989
6. [96%]	Likos, Christos N.	Effective interactions in soft condensed matter physics	
7. [96%]		Applications to soft condensed matter and biology	cop. 1994

Es wurden nun 37 Titel gefunden.

Wird nun ein Titel zur Volltitelanzeige aufgerufen, so die Klassifizierung in englischer Sprache angezeigt.

Title	Forces, growth, and form in soft condensed matter : at the interface between physics and biology / edited by A.T. Skjeltorp and A.V. Belushkin
Imprint	Dordrecht : Kluwer Academic Publishers, 2004

Subjects	SOFT CONDENSED MATTER : 538.96
	PATTERNS + PATTERN FORMATION (MOLECUL.SYST.,PHYSICS) : 539.21*5
	BIOPHYSICS + BIOPHYSICAL CHEMISTRY : 577.3

